

20 April 2024

Dr. Guillaume Souchay
Recommender
Peer Community In Ecology

Dear Dr. Souchay,

My co-authors and I thank you and the referees for helpful feedback on our manuscript originally titled, “Integrating multiple datasets to align biological and statistical populations for abundance estimation.” Per comments, we retitled this revision as “Using multiple datasets to account for misalignment between statistical and biological populations for abundance estimation.” In *the Peer Community In Ecology* database, this manuscript is identified as #640.

Here, we address the two main points identified by you, and below we detail responses to additional comments by you and the referees.

1. We understand your concern about our use of the terms “biological population” and “statistical population,” though we view it as more of a distinction than an opposition. Further, we think that the concept is important, and it is often ignored in the ecology literature, which is what prompted the angle of our manuscript in the first place. It is worth noting that many others before us have highlighted these two types of populations (selected citations included in the manuscript), but the terminology to describe them has differed. We can provide specific page numbers to these citations, if that would be helpful for your review of our manuscript. As an example, we offer the following excerpt from Krebs (1999, p. 115):

• **Population**

- (a) In **statistics**, the universe of items under study is called the population. I will distinguish this concept as the *statistical population*.
- (b) In **ecology**, a population is a defined group of organisms of one species living in a particular area at a particular time. I will distinguish this concept as the *biological population*.

Our point is that the concept is not new! In our manuscript, we revisit this age-old, fundamental distinction between a biological and statistical population, and we apply it to a common estimation problem with contemporary methods.

Nonetheless, we agree with you that how we introduced the terms was confusing and therefore, we made some changes to improve clarity. Notable changes are as follows:

- a. We replaced “mismatch” with “misalignment.” Lack of appropriate sampling is what leads to the misalignment, so we don’t view these shortfalls as being mutually exclusive. Still, we agree that mismatch could be confused with other uses of the term (e.g., related to effects of climate change), and therefore, we opted to use “misalignment” instead.
- b. We agree with you that one way of framing the question is to ask how representative a portion of a population is of the entire population. We simply are framing it in a different, but equally valid, way. In our revision, we added a sentence stating the problem as one of representativeness, which we hope clarifies our point for some readers that are more

familiar with your phrasing. Thanks for the suggestion!

- c. Your statement of, “Such ‘mismatches’ can occur in count sampling but it can also be real with other kind of surveys” is precisely why we introduced the conceptual distinction of these two populations so broadly and not related to sampling, per se. However, it seemed to be too broad based on comments by you and the referees, so we rewrote the first three paragraphs of the Introduction into two new paragraphs focused specifically on sampling constraints. We hope this revised introduction to the topic helps address reviewer concerns while also not losing sight of our point.
- d. To assist the reader with understanding the relationship between the biological and statistical populations, and the sampling framework, we added a schematic diagram (Figure 1).

2. We also appreciate concerns about the clarity of our methods, as the accompanying data files and code were not available to you or the referees. As such, it is difficult to know how to address this point satisfactorily, so please let us know if we have missed anything related to this main point. We made three primary changes related to the description of our methods.

First, we removed the word “integrated” from the manuscript, including the title, with one exception. While we agree with you and the referees that our model does not adhere to the definition of an “integrated model” by Kery and Royle (2020), it does fit the description provided by Zipkin et al. (2021). Kery and Royle (2020) state that an integrated model “has at least one parameter identical across data sets being combined,” which is not necessarily the case with our model. Conversely, Zipkin et al. (2021) wrote that an integrated model “includes multiple data sources within a single statistical modeling framework,” which is the case with our model. The primary reason we used data integration in our study was to ensure that error was propagated correctly through all component models, which we view as a major benefit to integrated modeling approaches. Thus, we included one sentence about this advantage of data integration in our manuscript. Otherwise, we removed all other references to “integrated model” to be responsive to the concerns of you and the referees.

Second, as we discussed via email, it is unfortunate that you and the referees did not have access to some (or all?) of our data files and code during your initial review. I apologize for any mistake on my part related to this unintentional blunder. To the best of my knowledge, I have remedied the situation, and the files are available to reviewers at <https://datadryad.org/stash/share/ePh6smm5C9W3zCmKC5V8EXiyT4bn9hbE5k7ZvXkV8H8>. We will finalize our data submission with Dryad after our manuscript is recommended, at which time the files will be available to the public at <https://doi.org/10.5061/dryad.0cfxpnw8m>. Our reason for not doing so previously (and now) is that, as I understand it, no changes to the files are permitted after they are finalized.

Third, we attempted to bolster the Methods section as much as possible without it becoming cumbersome and repetitive. We also reviewed our code to ensure that variables are defined and named consistently with the manuscript and that annotation is sufficiently descriptive. Lastly, we re-ran each model to ensure all the pathways and naming formats are correct. To avoid confusion, we removed Appendix 4, which was intended to be a generalization of our JAGS code, but it seemed to be problematic for the referees.

Again, we thank you and the referees for useful comments on our work and manuscript. We hope our responses outlined here and the revised manuscript sufficiently address your concerns. If you have questions, or problems accessing the data files or code, please let me know. I will be working in a remote part of Zambia from 1 April 2024 to 5 January 2025. While I should have regular email access, I also am expected to have shorts periods without it. If you need something right away, please contact my co-author, Paul Lukacs, at paul.lukacs@umontana.edu. I look forward to hearing from you soon!

Sincerely,

A handwritten signature in blue ink, appearing to read "Michelle Kissling", with a stylized flourish at the end.

Michelle Kissling

Krebs, C.J. 1999. Ecological methodology. Addison-Wesley Educational Publishers. Menlo Park, California, USA.

RESPONSES TO SPECIFIC COMMENTS

Recommender – Guillaume Souchay

Introduction

1. We addressed this comment in the main body of our response letter. One note here: oftentimes the statistical population is a portion of the biological population, but it does not need to be the case. The statistical population is simply what is sampled, whereas the biological population has demographic meaning. We urge you to consult with Thompson (2004), specifically pages 13, 136, 307, 309, and even the index, as well as the other citations in the manuscript.

Thompson, W. 2004. Sampling rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Washington, DC.

Methods

2. We only did one survey in 2009 because our engine failed, and it took several weeks to get a replacement part delivered to remote Alaska. We reworded the sentence to explain.

3. Fair point about timing of capture. However, we started to capture before the birds arrived, which was terribly unproductive (!), and we continued to capture until we lacked sufficient darkness for the night-lighting technique to be effective. Also, we examined this issue already, available in Kissling et al. (2016). The bottom line is that this species demonstrates boom-and-bust reproductive patterns among years, so we have not found that the timing of capture within a year makes any difference.

Kissling, M.L., P.M. Lukacs, S.B. Lewis, S.M. Gende, and J. Waite. 2016. Breeding decisions of a declining seabird, the Kittlitz's murrelet. *Marine Ornithology* 44:171–182.

Data analysis

4. Line 256. We addressed this comment in the main body of our response letter. We removed the phrase “integrated model” from our manuscript.

5. Line 270. Yes, most likely, though until we find a satellite transmitter that works for this species, we have no way of confirming it. Regardless, they make up a small portion of our sample (<2% locations).

6. Line 290. It was related to this study, but we published it separately many years ago, available in Lukacs et al. (2010). We saw no point in repeating that work here when we could simply reference it.

Lukacs, P.M., M.L. Kissling, M. Reid, S.M. Gende, and S.B. Lewis. 2010. Testing assumptions of distance sampling of a pelagic seabird. *Condor* 112(3):455–459.

7. Line 321. Thanks for this comment. We removed Appendix 4, as we think it caused more confusion than clarity. We also updated the annotation of our code to correspond to the manuscript, and we noted where deviations occurred.

8. Lines 401–403. We appreciate this comment, as we have another manuscript in review now that compliments this one. While this manuscript addresses within-season temporary emigration, the other manuscript addresses among-season temporary emigration. Your comment pertains to the latter problem, i.e., birds that arrive and decide not to breed or attempt to breed and fail early. As non-breeders, these birds do not need to remain in the study area, though some do. Thus, your comment here is beyond the scope of this paper and is addressed in our other paper, which uses the results of this study in an integrated population model (IPM). We referenced it in our revised manuscript as follows:

Kissling, M.L., P.M. Lukacs, S.M. Gende, G.W. Pendleton, and J.J. Felis. *In review*. Accounting for non-breeding when estimating population dynamics of a dispersed-nesting seabird, the Kittlitz's murrelet. *Ecosphere*.

Discussion

9. Lines 498–508. Yes, it was the same telemetry set. We agree with your comment about the usefulness of an IPM, and we did just that! In hindsight, it might have been best to publish the two manuscripts back-to-back in the same journal. It was too much information for one manuscript, so we divided it into two components: within-season temporary emigration due to bird movement and the dynamic ice habitat (this manuscript; sampling variance) and among-season temporary emigration due to non-breeding (the IPM; process variance). We can provide you with a draft of the IPM manuscript if that would be helpful.

10. Lines 595–598. Yes, the IPM was quite useful. However, a key component of the IPM is the results of our abundance model described in this manuscript.

Reviewer 1

Major comments:

1. As stated in our response letter, we tried to flush out the Methods section to help with clarity. We are not sure how to clarify some methods even further. For example, we estimated p_d using a standard distance sampling analysis, i.e. a half-normal detection function. Our manuscript is not about distance sampling, as this commonly used approach is well documented in the literature, including Kissling et al. (2007), which outlines our boat survey protocol in detail and is cited in the manuscript. The reviewer stated that the code in Appendix 4, which we removed per comments of other reviewers, was not sufficient to understand the analyses, yet the code includes the equations, which the reviewer would like to have seen. Therefore, we are confused as to how to address this comment.

Nonetheless, we added some language to the Methods section, clarified and numbered the equations in the manuscript, identified the equations in the JAGS code, and further annotated our code. We are not sure what more to do to address this comment, other than pages of equations or paste our code into the body of the manuscript.

2. As stated in our response letter, we removed the term “integrated model” from our manuscript.

3. The focus of the study was to account for within-season temporary emigration when estimating annual abundance. Because we could not sample the entire study area and the size of

the study area changed constantly due to ice floes, the statistical population, which is what was sampled, did not align with the biological population, which is what we wanted to know about. We purposefully used a single analytical framework to ensure that variability associated with each parameter was properly included in the resulting estimates. So, we agree with the reviewer and perhaps don't understand their concern.

Intuitively, precision of probability of presence increased as the window length increased, as the sample size increased (number of tagged birds and telemetry flights). We standardized the telemetry data by presenting results for the 3-day window only in the main body of the manuscript, though provided the results for all window lengths in the appendices for completeness. Again, we don't fully understand the reviewer's concern.

We agree with the reviewer that our results were sensitive to the number of boat surveys conducted in a year. This finding emphasized the importance of a second boat survey, as results of a single survey were overly precise and did not accurately represent the variability surrounding the abundance estimate. Certainly, we could have compared the statistical population abundance generated for boat surveys 1 and 2, though what would we gain? Isn't it better to use all the available information for each year to estimate abundance?

Minor comments:

1. Line 31. Done.
2. Line 35. Done.
3. Line 39. We appreciate the comment, but yes, we intend to use unbiased here.
4. Line 76. We changed the Introduction significantly and replaced "mismatch" with "misalignment." We hope these changes sufficiently address the reviewer's comment.
5. Line 112. We deleted "paradoxically."
6. Line 129. Yes, p_a can be brief (e.g., diving bird underwater) or long (e.g., animal in a burrow).
7. Line 161. We intended to use "unbiased" here.
8. Line 165. The other three objectives are explicit, whereas the fourth aim is more general. We left it as is.
9. Lines 176–192. We simplified the information describing our study area. We also added the label 'Malaspina Glacier' to Figure 2 (revision).
10. Line 196. It depends on the year, but within the 1-15 July timeframe. Sometimes, we had to wait for weather to improve in between surveys; other times, we had other tasks to accomplish (e.g., changing batteries at nest monitoring cameras); and still others, we encountered typical field problems like outboard engine issues. Our protocol is described in detail in Kissling et al. (2007), which is cited in the manuscript.

As an aside, in this publication, we recommended the current study, i.e., the manuscript in review at PCI Ecology, because of concerns about temporary emigration during our boat surveys.

11. Line 200. The transects shown in Figure 2 (revision) are our permanent transects that we aim to sample annually. However, because of ice, shallow water, etc., we rarely can complete a full set of transects for access reasons. We clarified this point in the text, per the reviewer's comment.

As for the effect of environmental conditions on detection probability, truncation distance, and possible double counting, we refer the reviewer to our protocol described in Kissling et al. (2007), which is cited in the manuscript. These points are valid, though they are part of standard distance sampling analysis, which is not the point of our paper. In short, we found that environmental covariates did not help explain variation in p_d (like group size), truncation distance varied because we applied a simple cutoff of 5% of all observations, and our transects are spaced at least 2 km apart, so we are not concerned about double counting.

12. Line 211. Yes, "in" Icy Bay refers to the actual bay, which comprises Main Bay, Taan Fjord, East Bay, and Upper Bay, and "near" Icy Bay refers to outside of it (e.g., Ocean stratum). To avoid confusion, we changed all of these phrases to be "in the Icy Bay study area" and referenced what is now Figure 2 in the revised manuscript.

13. Line 215. Yes, the frequencies were unique to each bird. How else would we track individuals? We included transmitter battery life (110 days). Also, at the end of the following paragraph (line 230 of the original manuscript), we refer the reader to several other publications that contain detailed information on tagging. Again, the point of this manuscript is about accounting for misalignment of the statistical and biological populations, not to assess our tagging effort, which has been done elsewhere in Open Access publications.

14. Line 223. We referred the reader to the study area as shown in Figure 2 (revision). We are not sure how to address some of the comments here, as many of them are related to aerial telemetry methods, not our study. Like the previous comment, at the end of this paragraph, we refer the reader to several other publications that describe our procedures for relocating tagged murrelets.

15. Line 270. These individuals are not part of the biological population, as we defined it. They likely were birds transiting through our study area, especially given that none of them were located again.

16. Line 298. Yes, we have a strict protocol for conducting surveys, which is outlined in Kissling et al. (2007), so the survey conditions are intentionally the same, or very close to the same.

17. Line 311. We added a line stating that p_k is indexed by strata in each year (not by survey). Also, this indexing is included in the first equation with the joint likelihood, and the fourth equation shows how p_k was calculated.

18. Line 312. No, p_p does not have strata but is estimated for each survey, as the indexing in the first and second equations show. We understand how the reviewer got confused, so we added a phrase for clarification.

19. Lines 321/330/345. The term y_p is the data. We didn't think it needed to be defined, as it is fairly standard notation. We added phrases to clarify it.

20. Line 328. We changed the indexing for each individual location from i to j in the code to be consistent with indexing in the manuscript. We also added a line leading into equation 2 that explains the relationship between p_p and beta. We did not consider individual covariates when modeling p_p , which we think is what the reviewer's comment was related to; all individuals were equally likely to be present (and absent) in the sampling frame.

21. Line 339. We retained the original notation, as the i refers to the survey in which group q was observed and therefore, this notation is consistent with the indexing of the data structure.

Regarding use of q as an index of sigma, when estimating detection function in distance sampling, the sampling unit is the group not the individual. So, use of q here is not referring to a group size effect, but instead the group itself. It is standard notation in distance sampling literature.

22. Line 388. We specified that λ refers to the geometric growth rate. Also, we changed λ to λ_g in equation 5 to distinguish it from the geometric growth rate, λ .

23. Line 403. Correct, we used 516 telemetry locations to estimate p_p in the 15-day window across all years.

24. Line 414. Fair point, though we do not have enough data to perform the analysis suggested. While the same telemetry survey data were used to estimate p_p in 2012, it was the only year where it was the case.

25. Line 422. Tidal stage is either incoming or outgoing, so yes, it is always binary and is related to the motion of the water. Because all readers may not be familiar with this phrase, we changed it to "tide direction." In Figure 4, the scale is linear so it could be displayed in similar to the other covariates. Also, the covariates on the x-axis are scaled.

26. Line 437. Done.

27. Line 477. Done. We added a phrase for clarification.

28. Line 481. No, p_p and boat survey effort are confounded.

29. Line 487. Done. We added a phrase for clarification.

30. Line 488. Done. We added 'temporal.'

31. Line 499. We added ‘generally’ to this sentence.

32. Line 563. We added ‘on murrelet movements.’

33. Line 571. In our study, we were not limited by the number of murrelets we could capture and tag, but instead by the number we could track. Unfortunately, we still lack a suitable satellite transmitter for this species, though that would dramatically improve our ability to account for temporary emigration during boat surveys. To answer the reviewer’s question, we do not know how many would need to be tagged, mostly because that is not our limitation currently.

34. Line 577. With satellite transmitters, tracking could occur beyond a single study area without additional effort.

35. Figure 1. We included a sentence in the caption noting that the gray-shaded area is land. The nests are located on the map as open circles (i.e., telemetry locations). We added a sentence to the Methods section stating that the outer limit of the Ocean state was constrained by aircraft gliding distance to shore, and it was the same boundary used in a multi-state survival analysis (Kissling et al. 2015b).

36. Figure 2. The boxplot format for c and d is the same as a and b, but the posterior distributions for c and d were narrower. We added ‘teal’ after posterior distribution in the caption.

37. Figure 3. We changed the axes. The reason we plotted standard errors instead of 95% credible intervals was because the credible intervals were much larger and required expanding the axis scales. It made the figure look out of proportion, unnecessarily. We agree that it usually is best to be consistent in presentation of results but thought the standard errors were more appropriate here (and in Figure 5).

38. Figure 4. We tried to add lines and/or ribbons (as suggested also by Reviewer #3) to this figure to show the relationship between the covariates and predicted p_p , but it was misleading. It looked as if we were trying to demonstrate a relationship when in fact there was no relationship. We retained the original figure in our revised manuscript, though changed the caption to emphasize the distribution of the predicted values (as opposed to the uncertainty of them).

39. Figure 5. We changed the axes and removed the word ‘mean’ from the caption. See reply to comment #37 above for explanation of use of standard errors.

40. Figure 6. Done.

41. Appendix 1. Done – thank you! In 2012, we experienced a lot of challenges with ice floes, which slowed us down tremendously. Therefore, the first boat survey on 9–10 July and the second boat survey on 11–12 July. Also, our aircraft was called for an emergency rescue on 9 July, which resulted in a cancelled telemetry flight, so we only had telemetry data for 11 July. Because we ultimately used a 3-day window to estimate abundance and trend, this data limitation was moot.

42. Appendix 3. Done.

Reviewer 2

Major comment:

1. As we stated in our response letter, we understand the reviewer's comment, though we disagree with it. We think that the conceptual distinction between a statistical and biological population is important and often ignored in ecological literature. Ultimately, we are simply talking about terminology, as this reviewer still has names for the "populations," i.e. "total population" and "statistical population."

Moreover, we found the reviewer's example of a "statistical population of breeders" to be an oversimplification of population dynamics and nature, in general. Even with breeders and non-breeders, some sort of misalignment occurs, as we demonstrate with our IPM (Kissling et al. *In review*). Our manuscript offers consistent terminology to refer to such a misalignment and one approach to accounting for it.

Minor comments:

1. Done. We opened the Abstract differently.

2. Done.

3. Line 63. We addressed this comment above. However, we want to point out that we do agree with the reviewer's point: the sampling frame defines the statistical population. We hope our revised Introduction satisfies this concern.

4. Line 87. In this context, a statistical driver relates to constraints of the sampling design.

5. Line 301. We clarified and added a citation here. When performing distance sampling analyses, a useful rule of thumb is to truncate about 5% of the data from the right-hand tail of the detection function.

6. Line 328. We added an equation to lead into equation 2, which we hope addresses this comment.

7. Appendix 4. We deleted the code from the manuscript.

Reviewer 3

Major comments:

1. We appreciate the reviewer's comment about presentation of the model, and we apologize for the confusion when reading our original manuscript. We can see how the confusion arose and hope that we resolved the issues by (1) numbering our equations so we could refer to them in the text and code; and (2) adding a short paragraph at the end of each section of "Data analysis" describing our modeling approach to clarify our workflow.

2. We bolstered annotation of our code, improved consistency in variable names, and added

numbers to the equations in the manuscript, which are referred to in the code. We thank the reviewer for identifying these shortcomings, as we think it is much better now.

3. This comment refers to the fact that the reviewer did not have access to the data files, which we resolved. Also, we removed Appendix 4. We thought the generalized code in Appendix 4 might be a simple reference for the reader, but it seems to have caused a lot of confusion, for which we apologize. We did not provide a GitHub link because of the instructions for submission on the PCI Ecology website. Instead, we used Dryad.

Minor comments:

1. Line 35. Per a comment from Reviewer #1, we removed ‘*r*’ and included ‘temporal trend’ instead.

2. Line 36. It refers to both.

3. Lines 64–66. We introduced the topic differently and hope it is clearer now.

4. Lines 76–83. We introduced the topic differently and hope it is clearer now. We also included an example.

5. Line 105. Done.

6. Line 107. As this sentence is a topic sentence, we left it as is. We discuss how some habitats are too complex in the paragraph.

7. Line 117. We added a phrase to link the thoughts together better. Thanks!

8. Line 120. We hope that our revised Introduction lays the foundation better. We left the sentence as is.

9. Line 132. Done.

10. Line 163. We added “...their variation across space and time.”

11. Lines 195+. As stated previously, we hope that this comment is resolved by having access to the data and code files.

12. Line 198. We only did one survey in 2009 because our engine failed, and it took several weeks to get a replacement part delivered to remote Alaska. We reworded the sentence to explain.

13. Line 214. We determined birds were more than 2 years old using plumage characteristics. We detail our capture, handling, tagging, and relocating methods in several other publications, which we refer the reader to on line 230 of the original manuscript.

14. Line 263. We purposefully describe them in the order of p_s , p_p , p_a , and p_d because they are

conditional on each other. In other words, a bird can't be detected if it is not available, present, and part of the statistical population. We left it as is.

15. Line 277. Done.

16. Lines 328–330. We added a new equation leading into equation 2 that we hope addressed this comment.

17. Lines 367–376. It is part of the data and code files on Dryad now.

18. Lines 380–381. In our state space model, we weighted the response variable by the inverse of its variance to give weight to the more precise estimates in our data. This practice is standard in regression, and we did not think it needed further explanation.

19. Line 393. Thank you for the comment.

20. Lines 399+. The reason we reported and plotted standard errors instead of 95% credible intervals for results in Figures 3 and 5 was because the credible intervals were much larger and required expanding the axis scales. It made the figures look out of proportion, unnecessarily. We agree that it usually is best to be consistent in presentation of results but thought the standard errors were more appropriate in these cases.

21. Lines 442–443. To manage runtime of the abundance models, we ran each year separately and therefore, we ran the state space model separately. We included a few sentences in the Methods section to clarify our workflow.

22. Lines 450–452. We changed the sentence to make it clearer. We agree with the reviewer that it was confusing!

23. Line 463. See response for *Major comment 1* for this reviewer.

24. Line 488. We retained 'trend' and deleted 'mean r_t ;' it was redundant.

25. Line 545. Done.

26. Line 623. Done. However, please note that the Dryad link here is the future permanent link. Reviewers must use the reviewer-only link:

<https://datadryad.org/stash/share/ePh6smm5C9W3zCmKC5V8EXiyT4bn9hbE5k7ZvXkV8H8>.

We believe this issue is where our data and code were lost during reviews of our manuscript.

27. We tried to add lines and/or ribbons (as suggested also by Reviewer #1) to this figure to show the relationship between the covariates and predicted p_p , but it was misleading. It looked as if we were trying to demonstrate a relationship when in fact there was no relationship. We retained the original figure in our revised manuscript, though changed the caption to emphasize the distribution of the predicted values (as opposed to the uncertainty of them).